The Coherence Engine, part deux

Collapsing Bayesian Inference, Free Energy, Gradient Descent, Markov Models, and Reinforcement Learning

Benjamin James

December 9, 2025

Abstract

Modern AI is driven by five dominant techniques, gradient descent, Bayesian inference, variational free-energy minimization, Markov models, and reinforcement learning. Although historically independent, these methods share a structural assumption; each optimizes a scalar objective inside a fixed representational space. Their successes arise when the world remains compatible with these frozen ontologies; their failures emerge whenever structural novelty is required. This paper reframes these techniques as constrained projections of a deeper coherence dynamic. When the representational space Ω is held rigid, coherence ascent collapses into the familiar behaviors of parameter tuning, posterior updating, or policy optimization. When Ω is allowed to evolve, the full Coherence Engine becomes visible. Adaptive model expansion, structural pruning, and curvature-driven transitions restore the system's ability to generate new abstractions, causal structures, and strategies. The future of AI depends not on improved optimization within predetermined models but on principled mechanisms for the evolution of models themselves. Coherence dynamics provides that mechanism, unifying inference, prediction, control, and representation under a single meta-law.

Keywords: coherence engine, ai foundations, structural evolution, bayesian inference, free energy principle, reinforcement learning, markov models, gradient descent, ontology change, adaptive systems

AI's Five Operational Axes

Modern AI is anchored by five methodological pillars that have come to define its practice: gradient-based optimization, Bayesian inference, variational free-energy minimization, Markovian structure, and reinforcement learning. These techniques emerged independently, shaped by distinct disciplinary pressures; calculus and cybernetics for gradient descent, probability theory for Bayes, information theory and physics for free-energy, stochastic processes for Markov models, and behavioral optimization for reinforcement learning. Their trajectories diverged historically, and within their respective traditions each appears conceptually self-sufficient. Yet when examined at the level of operational structure rather than historical genealogy, a striking convergence becomes visible; they all enact the same underlying move. Each optimizes a scalar function inside a fixed representational architecture.

This architectural fixity is rarely acknowledged because each technique internalizes its own boundaries as part of its mathematical identity. Gradient descent presupposes a predetermined architecture whose parameters can be adjusted but not reimagined. Bayesian inference assumes a closed hypothesis space and explores only the redistribution of credence within it. Free-energy minimization constrains the system to a generative model whose form cannot shift during inference. Markov processes encode a static state graph that structures what transitions are even conceivable. Reinforcement learning, despite its reputation for flexibility, takes its reward structure, action set, and environmental ontology as given. Across all five methods, one sees the same implicit assumption, *the system may optimize within its representational frame, but it may not transform the frame itself*.

This assumption is philosophically consequential. A theory that holds representational structure fixed cannot account for the very phenomena that most distinguish intelligence from mere optimization, the emergence of new categories, causal distinctions, abstractions, strategies, and dimensions of modeling. It cannot explain how systems adapt not only their beliefs or policies, but their ontologies, the internal structures through which they make the world intelligible. To optimize a function inside a fixed model is to behave like a well-tuned mechanism. To transform the model when coherence falters is to behave like something

closer to a learning organism, but modern AI overwhelmingly performs the former and has no internal principle for the latter.

These five operational axes did not fail to include structural evolution because of oversight. Rather, structural evolution lies outside the conceptual vocabulary of each method. Gradient descent cannot produce new dimensions; it can only scale or rotate the ones provided. Bayesian inference cannot invent hypotheses; it can only update weights. Free-energy systems cannot restructure generative models from within; they only refine latent variables. Markov processes cannot alter their state graphs; they merely populate them. Reinforcement learning cannot modify the ontology of actions, objectives, or environmental primitives. What these approaches optimize is the behavior of representations, not the evolution of representational structure.

This distinction matters now more than ever. AI has reached a point where further progress is constrained not by insufficient optimization power but by insufficient structural adaptability. Deep models can fit almost any function but cannot reorganize their abstractions when the world shifts outside their training distribution. Probabilistic models can refine uncertainty estimates but cannot recognize when their own hypothesis space is inadequate. RL agents can maximize returns but cannot identify that the reward structure itself encodes a mis-specified ontology. In every case, the limitation appears differently, but the underlying problem is identical. These systems lack a law of structural evolution.

The Coherence Engine, introduced in my prior paper, provides precisely this missing principle. It offers a unified meta-dynamic governing how systems preserve, refine, and transform their internal structures in response to coherence gradients. From this vantage, these five dominant AI techniques no longer appear as foundational theories but as restricted projections of coherence dynamics under severe constraints. Specifically, the constraint that the possibility space Ω remains fixed. When Ω is frozen, coherence ascent degenerates into local optimization inside a rigid frame; lift that constraint, and the deeper unity between these methods becomes visible.

This paper exposes that unity. The subsequent sections show how these operational axes reduce to a special case of coherence dynamics, what structural assumptions make that

reduction possible, and what forms of intelligence are lost when these assumptions are enforced. The goal is not to diminish the achievements of these methods but to situate them within a more general theory; one in which structural evolution is not an external intervention but a native operation of the system itself.

Gradient Optimization as Fixed-Architecture Coherence Ascent

Of all the techniques that define contemporary AI, gradient-based optimization is the most pervasive. It is the engine beneath almost every modern model of transformers, diffusion networks, graphical models, deep classifiers, generative architectures, reinforcement-learning policies, implicit models, and autoregressive decoders. Regardless of surface differences, each system traces its lineage back to the same operation, compute a loss, take its gradient, and update parameters. This uniformity is rarely acknowledged because it is masked by architectural proliferation. The universality of gradient descent is not a coincidence; it follows from a deeper constraint that modern AI takes for granted. Gradient descent thrives precisely because it presupposes a frozen representational architecture, a rigid Ω, inside which coherence ascent reduces to the familiar optimization of a scalar loss. From the perspective of the Coherence Engine, gradient optimization is the most extreme restriction one can impose on coherence dynamics. The loss function corresponds to a narrow instantiation of the coherence density C, usually tied to predictive fidelity, reconstruction accuracy, or reward. The parameters X are allowed to vary, but the structure that interprets those parameters, the architecture, hypothesis class, and inductive biases encoded in its wiring, is held fixed. Under these constraints, the general coherence update rule collapses into,

$$X_{t+1} = X_t - \eta \nabla L(X_t),$$

a purely local step on a static manifold. The representational ontology, what the system can express, infer, or conceptualize, is determined entirely by the architecture's initial specification. Coherence ascent becomes nothing more than loss descent inside a space whose structural boundaries are immovable.

This architectural rigidity is not accidental; it is foundational to gradient-based learning. For gradients to exist, the architecture must have a differentiable parameterization. For backpropagation to function, the computational graph must be predetermined. For optimization to succeed, the mapping from parameters to outcomes must be stable throughout training. In practice, this means that every deep learning system is defined by a static ontology. The network's layers, nonlinearities, attention patterns, and latent partitions specify all possible representational forms the model may ever discover. Training simply locates a high-fidelity configuration within that predetermined space.

Such rigidity becomes especially visible when neural models confront data or environments that violate their structural assumptions. Distribution shift exposes the limits of architectures tuned to one coherence regime but blind to others. Catastrophic forgetting reveals the absence of mechanisms for stable, multi-regime coherence maintenance. Out-of-class patterns, whether novel causal structures, new compositional forms, unanticipated strategies, or unseen abstractions, cannot be represented unless the architecture already contains the latent structure required to encode them. Even self-supervised foundation models, despite their scale, inherit this limitation. They interpolate extraordinarily well within the manifold they were trained on, but they rarely generate new representational axes when the world deviates from learned regularities.

These failures are not pathologies of gradient descent; they are expressions of its philosophical commitment. Gradient-based learning assumes that all structural degrees of freedom must be specified a priori. The system may refine its beliefs, but it may not alter its own conceptual vocabulary. It may adjust weights, but not the dimensionality or topology of the space those weights inhabit. This is the essence of its frozen-ontology problem. The model's expressive structures are held constant even when coherence demands evolution.

The consequences are systemic. Deep networks often respond to structural mismatch by saturating their gradients, stalling in low-coherence regimes without any means of expanding or restructuring Ω. When ∇L collapses, so does learning; the model cannot take a meaningful step because there is no direction within the current architecture that increases coherence. As a result, optimization terminates not in genuine understanding but in

representational exhaustion. The system has reached the limits of what its fixed form allows it to grasp.

In contrast, coherence dynamics do not privilege architectural rigidity. When ∇C stagnates inside a fixed Ω, the Coherence Engine interprets that stagnation not as convergence but as an indication that the representational space is insufficient. The appropriate move is not a smaller learning rate, a new regularizer, or a different initialization; it is a structural adjustment, expansion or reconfiguration of Ω itself. This may correspond to adding new latent dimensions, introducing new structural motifs, reshaping dependency patterns, or reparameterizing the entire representational manifold. Coherence ascent continues not by optimizing within a given architecture, but by changing the architecture so another coherence gradient becomes accessible.

In this light, gradient descent appears not as a fundamental learning principle but as a degenerate limit of coherence dynamics under extreme constraints. It functions only when the world is sufficiently compatible with the model's fixed ontology. When that compatibility breaks, when new coherence structures must be represented, gradient systems reveal their philosophical inflexibility. They can polish what already exists, but they cannot invent the structural substrate required for deeper understanding.

Recognizing this limitation is not an indictment of gradient-based learning; it is an acknowledgment of the conceptual boundary that defines it. The purpose of situating gradient descent within coherence dynamics is to make explicit what the method itself cannot do, generate new representational axes when coherence falters. The Coherence Engine fills precisely this gap, not by improving optimization within the model, but by providing the law that determines how the model itself must evolve.

Bayesian Inference and Variational Free-Energy Methods as Closed Ω Inference

Bayesian inference and variational free-energy methods are widely regarded as the most principled approaches to uncertainty, reasoning, and information integration in modern AI. They provide rigor where gradient systems offer heuristics, and interpretability where deep neural models offer opacity. Yet these methods share an assumption that is rarely

articulated in discussions of their power, both operate within a static possibility space, a fixed Ω, whose structure is presupposed by the modeler. Every posterior, every variational distribution, and every generative model is embedded within a representational framework that is itself immune to revision. As with gradient descent, but expressed in a different vocabulary, Bayesian and free-energy systems optimize internal quantities while taking their ontological scaffolding as given.

Bayesian inference begins with a hypothesis set $\Omega = \{h_1, h_2, \dots\}$, a prior distribution over it, and a likelihood model. The update rule is well known, posterior probability mass is redistributed toward hypotheses that better predict observed data. In coherence-engine terms, this corresponds exactly to coherence reweighting within a fixed Ω. The coherence density C collapses to a log-likelihood term plus (optionally) a complexity penalty, and the update rule becomes,

$$P(h \mid d) \propto P(d \mid h)P(h),$$

a pure redistribution of weight on a static hypothesis lattice. Everything depends on Ω, the system cannot add a new hypothesis, redefine an old one, or restructure the relations between them. Bayesian inference is therefore a special case of coherence ascent under the strictest constraint. Ω is permanently frozen, and the only permissible coherence adjustment is the shift of credence among predefined structures.

This rigidity becomes evident in Bayesian neural networks, which are often presented as the gold standard for principled uncertainty. Even the most sophisticated Bayesian neural networks operate within a hypothesis space determined entirely by the network architecture, prior distributions, and parameterization chosen at design time. While weight distributions may broaden or tighten, the architecture's representational ontology, its dimensionality, layer structure, and functional forms, remains fixed. The posterior accounts for uncertainty over configurations of a pre-established model, not for uncertainty over the model's structure itself. Changes in the world that require new representational dimensions, or new causal relations, cannot be accommodated except by manual architecture modification. Bayesian inference does not generalize to new ontologies; it only reallocates belief within old ones.

Variational inference and the Free Energy Principle (FEP) share this same structural constraint, though expressed through different mathematics. Variational inference replaces exact Bayesian posteriors with tractable approximations, selecting the $q(s)$ that minimizes a divergence from the true posterior. The variational free energy F functions as a coherence measure. In fact, negative free energy is a form of coherence density,

$$C = -F = \mathbb{E}_{q(s)}[\log p(o, s)] - \mathbb{E}_{q(s)}[\log q(s)].$$

This formal equivalence makes variational inference a direct instantiation of coherence ascent under the choice $C = -F$. However, the structure of $q(s)$, the structure of the generative model $p(o, s)$, and the parameterization of the latent space are all assumed fixed. The inference procedure adapts beliefs within this space, but the space itself cannot be reshaped by the dynamics. Even in hierarchical generative models, such as those seen in deep variational autoencoders, the hierarchy itself is static; its form chosen beforehand by human designers. The system infers latent causes, but it cannot infer the need for new representational motifs. It cannot declare that its generative model is missing a variable, missing a dependency, or missing an entire dimension of causal structure.

This constraint is even clearer in the Free Energy Principle's cognitive interpretation. FEP proposes that biological systems maintain coherence by minimizing variational free energy, effectively predicting sensory data through an internal generative model, but the generative model is taken as a given structure. Perception is cast as the adjustment of beliefs within a model, not the adjustment of the model itself. FEP describes inference over a fixed ontology; it does not describe ontological evolution. Innovation, conceptual restructuring, and representational emergence lie outside its formal machinery unless explicitly engineered.

Modern AI applications inherit these constraints. VAEs learn powerful latent representations, but the form of the latent space, its dimensionality, distribution family, and factorization, cannot shift in response to new patterns. Hierarchical generative models infer hidden causal factors only to the extent that the architecture already contains slots in which those factors can be expressed. Bayesian hierarchical models can infer hyperparameters, but they cannot infer that the underlying hypothesis space is malformed. In each case, the

system excels when the world lies within the manifold encoded by Ω and fails silently when the world departs from it.

The key limitation is this, *Bayesian and free-energy methods do not contain a rule for generating new hypotheses***.** They allocate coherence among predefined structures, but they lack a mechanism for revising the structures themselves. Their inference is epistemically flexible but ontologically rigid. When the world leaves the model class, when new causal relations, new temporal patterns, or new abstractions become necessary, the posterior does not scream in protest; it merely distributes probability mass across the closest available approximations. The inference procedure saturates, not because it has succeeded, but because the representational frame lacks the degrees of freedom needed for further coherence ascent.

This mismatch between inference and world dynamics becomes acute in environments that undergo structural change. A Bayesian neural network trained on iid data cannot represent a shift to non-iid structure unless that possibility was encoded in its initial architecture. A VAE cannot discover a new latent factor unless the modeler pre-allocated a latent dimension where that factor could reside. An FEP agent cannot restructure its generative model in response to conceptual novelty; it can only adjust posterior beliefs inside the model it already has. These methods are powerful precisely when the world conforms to their representational scaffolding, and brittle when it does not.

Coherence dynamics exposes the root of this brittleness. Bayesian and variational systems are coherent only within the closure of Ω. When ∇C collapses, when no reweighting or variational adjustment increases coherence, Bayes and FEP interpret this as convergence. The Coherence Engine interprets it as a signal that Ω is inadequate. The appropriate response is not posterior stabilization but structural expansion, a modification of Ω itself. The engine allows new hypotheses to emerge, new dimensions to form, and new generative structures to be proposed and evaluated. Coherence ascent continues by transforming the representational ontology, not merely refining beliefs within it.

In this light, Bayesian and free-energy methods reveal themselves as highly disciplined but structurally impoverished slices of a broader coherence dynamic. They succeed by fixing Ω

and perfecting inference over it, but inference over a fixed ontology is not enough when the world routinely demands new ontological categories. The limitation is not in their mathematics but in their metaphysics. They treat representational structure as given, not as something subject to evolutionary pressure. Coherence dynamics fills this void, providing the principle that Bayes and FEP lack, the principle that governs the evolution of structure itself.

Markov Models, Reinforcement Learning, & Coherence in Constrained State Spaces

Markov models and reinforcement learning (RL) occupy an unusual position in AI. They are treated as distinct methodological families, one probabilistic and generative, the other decision-theoretic and evaluative, but both rely on the same underlying structural assumption. The state space is fixed, transition structure is fixed, and coherence flows only along edges prespecified by the modeler. Whether one frames the system as an HMM, an MDP, a simulator, or a diffusion chain, the ontology of states and transitions is given, not learned. The system may refine probabilities or policies, but the architecture of possible worlds remains immovable.

The Markov assumption itself encodes this rigidity. A Markov process begins with a predefined set of states Ω and a transition kernel $P(s' \mid s)$ that governs how coherence, interpreted as probability mass, moves from one state to another. The agent or model does not question whether the state graph is adequate; it simply redistributes belief across transitions that were already deemed relevant. In coherence-engine language, a Markov chain is what remains when Ω is frozen at the level of states and transition relations, C reduces to local transition coherence, and no structural learning is permitted.

Hidden Markov Models (HMMs) exemplify this. They are widely used in speech recognition, time-series modeling, genomics, and sequential inference. However, their expressive power derives entirely from the ontology their designer selects, the number of hidden states, emission structure, and conditional independence assumptions. Once specified, the model may update transition probabilities, but it cannot add new states or alter the topology of its latent space. If the world introduces a new mode, a novel causal dependency, or an

unforeseen temporal pattern, the HMM cannot restructure itself to accommodate it. Instead, it redistributes probability mass across the existing graph, approximating novelty as noise rather than as an invitation to revise the ontology.

Diffusion models, despite their contemporary prominence, operate under the same Markov constraint. Their forward and reverse processes assume a fixed-dimensional space in which Gaussian transitions smooth and then recover structure. While the learned denoising network is expressive, the Markov kernel and the latent trajectory space are architecturally predetermined. These systems excel at refining structure inside a fixed coherence manifold but cannot generate new representational axes when confronted with fundamentally novel input regimes. The Markov assumption remains a structural boundary even when embedded inside deep neural architectures.

Reinforcement learning inherits this same rigidity but expresses it through a decision-theoretic vocabulary. An RL agent operates inside an MDP. It observes states from Ω, selects actions from a predefined action set A, experiences transitions according to $P(s' \mid s, a)$, and maximizes a scalar reward function $R(s, a)$. In this setting, coherence reduces to expected return, a one-dimensional function that compresses all structural considerations into a single scalar. Under the Coherence Engine's lens, RL appears as a coherence projection where $C = \mathbb{E}[\sum_t \gamma^t R_t]$ and Ω consists only of the state-action structure designated at design time.

This architecture induces three layers of ontological rigidity:

1. A fixed state graph.

The agent must interpret the world through a predefined set of states. It cannot invent new states to represent new causal or conceptual distinctions. Even in continuous spaces, the representational basis is specified by the function approximator's architecture.

2. A fixed action set.

The agent cannot conceptualize actions outside the provided set, cannot abstract actions, and cannot generate new strategic operators. Exploratory behavior searches within the given space, never beyond it.

3. A fixed reward ontology.

Rewards encode the system's normative structure. They determine what counts as coherence, value, or success. Yet RL has no mechanism to question or revise this ontology; it can only optimize relative to the one provided.

These constraints explain why RL agents struggle to generalize abstractly or creatively. They cannot discover new forms of coherence, new ways of engaging the environment, because those forms lie outside Ω. They optimize only inside the manifold defined for them. When they face tasks that demand novel causal representations or new action abstractions, their failure mode is not partial degradation; it is representational paralysis. The coherence gradient collapses not because the task is solved, but because Ω contains no path through which coherence could increase.

This analysis reveals a structural unity between probabilistic sequential models and decision-theoretic ones. HMMs and MDPs differ in purpose, prediction vs. control, but they share the same metaphysical constraint. The world is reduced to a set of predefined states, and all adaptation operates by adjusting parameters within that state space. In both systems, coherence flows through fixed channels. The agent does not infer new states; it merely traverses those it was given. It does not restructure the transition graph; it simply learns probabilities over the existing edges. Learning is confined to intensities, not architectures.

Viewed through the Coherence Engine, these frameworks are one-dimensional coherence projections. They compress the multi-scale, multi-modal coherence landscape into a narrow scalar function (likelihood, return, or transition fidelity), and they bind that function to a rigid representational topology. As a consequence, they can perform local adaptation, parameter reweighting, policy refinement, or transition estimation, but cannot perform structural adaptation, transformation of state space, addition of new causal primitives, emergence of new action repertoires, or redefinition of the reward ontology.

When the world remains inside the Markov or MDP manifold, these methods excel. They provide clear, efficient, mathematically grounded tools for reasoning about uncertainty and action, but when the world leaves that manifold, when new structure appears, when causal

assumptions shift, and when novel forms of coherence become relevant, they fail by design. They can only represent new phenomena by deforming an inadequate ontology, never by revising it.

The Coherence Engine exposes why a Markov model or RL agent interprets coherence solely through the structural channels granted at initialization. Its ontology is external to learning. True generalization, abstraction, and creativity require Ω to evolve. Markov models and RL systems cannot evolve Ω from within; coherence dynamics can.

Convergence, Structural Constraints, and Collapsed Coherence

The five dominant techniques of modern AI, gradient descent, Bayesian inference, variational free-energy minimization, Markov models, and reinforcement learning, appear diverse when viewed from within their historical disciplines. Their equations differ, their motivations differ, and their conceptual vocabularies differ. Yet when examined structurally, they reveal a deep uniformity. Each assumes that the representational space Ω is fixed, and each performs a constrained ascent of coherence within that frozen space. Their differences are surface manifestations of this single architectural choice.

Fixing Ω has two immediate consequences. First, it collapses coherence into a narrow function, typically a scalar quantity that captures only one dimension of the system's interaction with the world, such as loss, likelihood, free energy, transition fidelity, or expected return. Second, it prevents the system from responding to coherence stagnation by expanding or restructuring its ontology. When ∇C approaches zero inside a fixed Ω, the system must treat this as convergence, because no representational move exists that would reopen the coherence gradient. The model cannot act on the insight that its representational frame is inadequate.

This structural constraint explains the characteristic failure modes of each technique. In gradient descent, stagnation appears as a vanishing gradient. The landscape offers no direction for further improvement within the architecture's existing degrees of freedom, even when a true coherence landscape lies beyond that architecture. In Bayesian inference, stagnation appears as a posterior concentrated on the least-wrong hypothesis, not because

the model is correct, but because Ω contains no hypotheses of the right type. In variational methods and free-energy minimization, stagnation appears as a variational optimum reached under a misspecified generative model, where the system refines beliefs inside an ontology that itself requires revision. In Markov models, stagnation appears as probability mass trapped in an inadequate state space, unable to represent novel regimes except by misinterpreting them as stochastic perturbations. In reinforcement learning, stagnation appears as policy convergence relative to a reward ontology that fails to express the structure of the task; the agent optimizes a coherence proxy that does not encode the true coherence landscape.

In each case, stagnation is not a signal of success but an artifact of architectural foreclosure. The model has reached the limit of what it can express, not the limit of what the world affords. These are not five unrelated forms of brittleness; they are five manifestations of the same underlying rigidity.

This unity becomes clearer when one imagines the coherence landscape geometrically. If coherence is a multi-dimensional field defined over the space of possible representational structures, then each classical technique operates on a projection of that field onto a lower-dimensional surface. Gradient descent follows the projection of ∇C onto the parameter manifold of a fixed architecture. Bayesian inference follows the projection of ∇C onto a discrete hypothesis lattice. Variational methods follow ∇C within a restricted family of variational distributions. Markov models follow ∇C over a predefined transition graph. RL follows ∇C along action policies that seek to maximize a one-dimensional reward scalar. Each technique thus traverses a thin slice of coherence space, and each converges when that slice offers no additional ascent direction, regardless of whether coherence remains available in the unreached dimensions.

This is why these five methods converge conceptually despite their superficial differences. They share not a common formula but a common absence, the inability to treat representational structure as something that can change in response to coherence gradients. Their mathematical forms differ because each chooses a different projection of

coherence onto a narrow function, but their philosophical core is identical. Each is a mechanism for optimizing within structures, not a mechanism for evolving structures.

Recognizing this shared constraint reveals why a unifying meta-dynamic is needed. Without a principle for Ω's evolution, AI systems will continue to exhibit local competence and global brittleness. Impressive performance inside the representational regimes they are designed for, but incoherence whenever the world shifts beyond those regimes. The Coherence Engine supplies the missing principle. When ∇C collapses inside a fixed Ω, this is not convergence but a signal that Ω must evolve. Structural stagnation is not an endpoint but the beginning of ontology revision.

By making the boundary shared across these five techniques explicit, this section dissolves their apparent independence and prepares the ground for my final synthesis, how the Coherence Engine restores the full coherence gradient that each method approximates in isolation.

The Coherence Engine as the Meta-Law Behind the Five Techniques

The previous sections have shown that the five dominant techniques in modern AI, gradient descent, Bayesian inference, variational free-energy minimization, Markov models, and reinforcement learning, are not foundational laws of learning but constrained projections of a deeper process. Each is what coherence ascent looks like when the possibility space Ω is frozen, when structural degrees of freedom are held fixed, and when coherence must flow along channels determined entirely by the modeler. Their power derives from their clarity within those constraints; their brittleness derives from the constraints themselves. To understand what lies beyond these techniques, one must remove the constraint on Ω and observe the dynamics that remain. That unconstrained dynamic is the Coherence Engine.

Removing Ω's rigidity restores the full geometry of coherence dynamics. Instead of performing ascent on a surface of fixed dimensionality, the system gains access to transitions that modify the space itself. When ∇C saturates locally, not because the world has been fully captured, but because the model lacks the representational capacity to capture it, the engine activates the mechanisms of adaptive Ω. These include model-class

expansion, structural innovation, and the formation of new representational dimensions. The system does not interpret stalled gradients as convergence; it interprets them as evidence that the ontology should evolve. Coherence ascent is no longer constrained to improve beliefs inside an inherited structure but now operates on the structure itself.

Adaptive Ω brings with it the capacity for structural pruning. Where expansion introduces new representational possibilities, pruning eliminates incoherent or low-curvature structures that no longer contribute to coherence. Unlike regularization in gradient systems or Occam penalties in Bayesian ones, pruning is not an external bias but an intrinsic consequence of coherence dynamics. Structures that contribute little to coherence or introduce persistent leakage decay naturally under the engine's dynamics. Pruning and expansion operate in tandem, producing a representational landscape that can grow in complexity when coherence demands it and contract when coherence would otherwise dissipate. This interplay gives the system a means of maintaining integrity across scales, expanding to capture new patterns and compressing to avoid incoherence.

The engine also introduces curvature-driven transitions, which govern when the system should shift representational regimes entirely. Whereas gradient descent only follows local slopes, coherence dynamics is sensitive to the second-order geometry of coherence fields. High curvature indicates that a region is structurally rich; low curvature indicates representational exhaustion. Curvature-driven transitions allow the system to reconfigure itself nonlocally, jumping to new structural configurations when the coherence landscape supports such a move. This capability is absent from all five classical techniques, each of which remains confined to continuous updates within a predetermined topology.

Viewed from this perspective, the classical techniques reappear as special cases when the engine's variables are selectively frozen. If Ω is fixed and C is chosen as a loss function, the system reduces to gradient descent. If Ω is fixed and C reduces to log-likelihood minus complexity, Bayes reemerges. If C becomes negative variational free energy and structural innovation is disallowed, FEP appears. Fix Ω as a transition graph and constrain coherence to transition fidelity, and Markov processes result. Fix Ω as an MDP and use a scalar C, and reinforcement learning emerges. These are not separate laws but constrained shadows of

coherence dynamics, each capturing one facet of the unified process when particular degrees of freedom are suppressed.

Next-generation AI cannot rely on the five classical techniques alone because the environments in which AI systems operate increasingly demand structural evolution. Abstraction requires expanding Ω to create representational forms that did not previously exist. Generalization requires recognizing when the current ontology cannot capture a novel regime and adjusting accordingly. Creativity requires exploring coherence-increasing structures outside the inherited manifold. Causal reasoning requires inferring not only parameter values but the structure of the causal graph itself, a quintessential Ω-modification problem. Self-modifying architectures require internal principles for structural revision, not external intervention by designers. These capacities cannot be approximated by optimizing inside fixed ontologies; they require a dynamic that governs the evolution of ontologies.

Many of the highest-performing AI systems today succeed not because they are structurally adaptive, but because the world is conveniently compatible with their frozen architectures. Large language models, diffusion systems, and reinforcement-learning policies all operate astonishingly well within their respective training manifolds, but they remain vulnerable to shifts in causal structure, distributional regime, or representational demand. Their failures are predictable once one recognizes that they lack a mechanism for coherence-guided structural innovation. The next step in AI development is not to refine these methods further, but to transcend the architectural rigidity at their core.

The Coherence Engine provides the meta-law necessary for this transition. It unifies the five operational axes of AI not by merging their formalisms but by revealing them as constraint-bound instantiations of a single coherence dynamic. Once Ω is allowed to evolve, the artificial distinction between optimization, inference, prediction, control, and generation dissolves. They become different aspects of coherence ascent across different structural regimes. Abstraction, generalization, creativity, causality, and architectural self-modification become natural consequences of the same underlying law. The evolution of models replaces the static optimization of models. With this shift, AI begins to behave not

like a mechanism tuned to a structure, but like a system capable of transforming the structure itself.

Optimization Inside Structures vs. Evolution of Structures

Across its many formulations, modern AI has perfected the art of optimization inside fixed representational structures. Gradient descent tunes parameters, Bayes redistributes belief, free-energy minimization refines posterior approximations, Markov models propagate coherence along predefined transitions, and reinforcement learning ascends expected return within a predetermined state–action ontology. These methods differ in language but not in assumption. Each presumes that the model's underlying structure is given, stable, and conceptually complete. Their successes arise precisely because the environments in which they operate, their datasets, simulators, and curated tasks, conform closely enough to their frozen ontologies for meaningful coherence ascent to occur.

The limitation becomes visible when the world does not cooperate. When causal regimes shift, when new structural patterns arise, or when abstractions must be invented rather than inherited, none of the five pillars can respond from within. They converge prematurely because their coherence gradients collapse prematurely, not in virtue of mastery, but because Ω contains no further representational dimensions in which coherence could increase. These systems behave as though the only meaningful intelligence is optimization within a structure, not the evolution of the structure itself.

The central realization of this paper is that the future of AI does not hinge on improving optimization inside fixed models. It hinges on developing systems whose representational structures evolve in response to coherence demands. The five classical techniques succeed only because the world is locally compatible with the ontologies they encode. The Coherence Engine, in contrast, provides the meta-law that governs how those ontologies must shift when the world no longer fits. It offers the rule not merely for adjusting beliefs or policies, but for revising the structural vocabulary through which belief and policy become meaningful.

The implications extend far beyond AI engineering. In physics, coherence dynamics reframes how laws emerge from structural regularities rather than residing in fixed ontologies. In cognition, it clarifies how minds generate new concepts rather than merely updating old ones. In evolutionary systems, it describes how new strategies and niches emerge rather than being enumerated in advance. For open-ended artificial intelligence, it supplies what has been missing since the field's inception, a principled mechanism for representational evolution.

Optimization shapes behavior inside structures; coherence dynamics shapes the structures themselves. This distinction marks the boundary between engineered intelligence and adaptive intelligence, and the path forward lies firmly on the latter side.

References

James, B. (2025). The Coherence Engine. PhilPapers. https://philpapers.org/rec/JAMTCE